

NETWORK-BASED REMOTE DATA STORAGE SYSTEM HAVING MULTIPLE ACCESS INTERFACES

This application is a continuation of U.S. Application Number 09/501,039, filed on February 9, 2000 the entire content of which is incorporated herein by reference.

TECHNICAL FIELD

This invention relates to storage of digital data, and more particularly to a data storage system that provides access to a remote storage area.

BACKGROUND

The Internet is a worldwide collection of networks that spans over 100 countries and connects millions of computers. Reports indicate that the Internet is growing faster than all preceding information technologies including radio and television. Remote data storage over the Internet is one of the fastest growing facets of the Internet. Numerous companies provide a wide range of an Internet data storage services for remotely storing and managing data files. Many of these services, however, have limited modes for accessing the data files. For example, many require that a user load proprietary software on his computer in order to communicate data files to the remote storage. Other services only support access via a web browser or similar communication utility.

SUMMARY

The present invention is directed to a data storage system and method capable of providing a variety of access methods. In one embodiment, a data storage system includes a plurality of storage servers coupled to a plurality of storage devices. The storage devices provide a plurality of virtual storage areas and each virtual storage area is assigned to a user. According to the invention, the assigned user can easily access the assigned virtual storage area using a client computer communicatively coupled to the storage servers via a global computer network such as the Internet. More specifically, in order to store data files within the remote virtual storage area, the user sends an electronic mail message that includes the data file as well as user information and target data file information. The storage server parses the electronic mail message and stores the data file within the storage area according

to the target data file information. In addition, the user can request one or more data files from the storage area and electronically mail the data files to the user. In this manner, authorized users can easily store data files to or retrieve data files from his or her remote storage area from anywhere in the world via a global computer network such as the Internet or a private wide-area network. The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

DESCRIPTION OF DRAWINGS

Figure 1 is a diagram showing an exemplary hardware and operating environment of a suitable computer for use with embodiments of the invention;

Figure 2 illustrates one embodiment of a computing system in which a storage server provides seamless access to remote storage areas;

Figure 3 illustrates a user interface displayed by one embodiment of an operating system for directly accessing a remote storage area using standard file management routines provided by the operating system;

Figure 4 illustrates a user interface displayed by another embodiment of an operating system for directly accessing a remote storage area;

Figure 5 illustrates an example operating system script written in AppleScript for the Macintosh® operating system that automatically copies a file from a local directory to a similarly named remote directory located within a storage area;

Figure 6 illustrates a window displayed by a conventional web browser when a user accesses a virtual storage area;

Figure 7 illustrates a window displayed by a conventional electronic mail software application when a user mails one or more data files to a corresponding storage area;

Figure 8 is a flow chart 800 providing a high-level overview of one mode of operation in which storage network 220 allocates storage areas.

DETAILED DESCRIPTION

In the following detail description, references are made to the accompanying drawings that illustrate specific embodiments in which the invention may be practiced. Electrical, mechanical, programmatic and structural changes may be made to the
 5 embodiments without departing from the spirit and scope of the present invention. The following detailed description is, therefore, not to be taken in a limiting sense and the scope of the present invention is defined by the claims.

The detailed description is divided into three sections. The first section provides definitions for various terms used throughout the detailed description. The second section
 10 describes an exemplary hardware and operating environment in conjunction with which embodiments of the invention can be practiced. Finally, the third section describes various systems, methods, potential benefits and alternative embodiments of the invention.

Definitions

15 **File Transport Protocol (FTP)** - A networking protocol specifically for transporting files from one computer on the network to another.

HyperText Markup Language (HTML) - an authoring language that defines the syntax and semantics used to create documents on the World Wide Web.

20 **Internet** - a worldwide collection of networks that spans hundreds of countries and connects millions of computers.

Internet Protocol (IP) - A low level communications protocol that specifies the format of individual communication packets and an addressing scheme.

Redundant Array of Independent Disks (RAID) - A high-volume storage device having multiple storage drives and fault recovery procedures.

25 **Server Message Block (SMB)** - a message format used by to share files, directories and devices.

Transmission Control Protocol (TCP) - a higher-level network communication protocol that establishes a virtual connection between a destination and a source.

30 **Transmission Control Protocol/Internet Protocol(TCP/IP)** - the suite of network communications protocols including both TCP and IP.

Web Distributed Authoring and Versioning (WebDAV) is a set of extensions to the

Hypertext Transfer Protocol (HTTP) that seeks to make the World Wide Web a collaborative, writeable medium.

World Wide Web (WWW) - A system of Internet servers that support documents specially formatted in the HTML language that supports links to other documents, as well as graphics, audio, and video files.

An Exemplary Operating Environment

Figure 1 illustrates a computer 100 suitable for supporting the operation of an embodiment of the present invention. As shown in Figure 1, the computer 100 includes a processor 112 that in one embodiment belongs to the PENTIUM[®] family of microprocessors manufactured by the Intel Corporation of Santa Clara, California. However, it should be understood that the invention can be implemented on computers based upon other microprocessors, such as the MIPS[®] family of microprocessors from the Silicon Graphics Corporation, the POWERPC[®] family of microprocessors from both the Motorola Corporation and the IBM Corporation, the PRECISION ARCHITECTURE[®] family of microprocessors from the Hewlett-Packard Company, the SPARC[®] family of microprocessors from the Sun Microsystems Corporation, or the ALPHA[®] family of microprocessors from the Compaq Computer Corporation. Computer 100 represents any server, personal computer, laptop or even a battery-powered, pocket-sized, mobile computer known as a hand-held PC or personal digital assistant (PDA).

Computer 100 includes system memory 113 (including read only memory (ROM) 114 and random access memory (RAM) 115), which is connected to the processor 112 by a system data/address bus 116. ROM 114 represents any device that is primarily read-only including electrically erasable programmable read-only memory (EEPROM), flash memory, etc. RAM 115 represents any random access memory such as Synchronous Dynamic Random Access Memory.

Within the computer 100, input/output bus 118 is connected to the data/address bus 116 via bus controller 119. In one embodiment, input/output bus 118 is implemented as a standard Peripheral Component Interconnect (PCI) bus. The bus controller 119 examines all signals from the processor 112 to route the signals to the appropriate bus. Signals between the processor 112 and the system memory 113 are merely passed through the bus controller

119. However, signals from the processor 112 intended for devices other than system memory 113 are routed onto the input/output bus 118.

Various devices are connected to the input/output bus 118 including hard disk drive 120, floppy drive 121 that is used to read floppy disk 151, and optical drive 122, such as a
5 CD-ROM drive that is used to read an optical disk 152. The video display 124 or other kind of display device is connected to the input/output bus 118 via a video adapter 125.

A user enters commands and information into the computer 100 by using a keyboard 40 and/or pointing device, such as a mouse 42, which are connected to bus 118 via input/output ports 128. Other types of pointing devices (not shown in Figure 1) include track
10 pads, track balls, joy sticks, data gloves, head trackers, and other devices suitable for positioning a cursor on the video display 124.

As shown in Figure 1, the computer 100 also includes a modem 129. Although illustrated in Figure 1 as external to the computer 100, those of ordinary skill in the art will quickly recognize that the modem 129 may also be internal to the computer 100. The modem
15 129 is typically used to communicate over wide area networks (not shown), such as the global Internet. Modem 129 may be connected to a network using either a wired or wireless connection.

Software applications 136 and data are typically stored via one of the memory storage devices, which may include the hard disk 120, floppy disk 151, CD-ROM 152 and are copied
20 to RAM 115 for execution. In one embodiment, however, software applications 136 are stored in ROM 114 and are copied to RAM 115 for execution or are executed directly from ROM 114.

In general, the operating system 135 executes software applications 136 and carries out instructions issued by the user. For example, when the user wants to load a software
25 application 136, the operating system 135 interprets the instruction and causes the processor 112 to load software application 136 into RAM 115 from either the hard disk 120 or the optical disk 152. Once software application 136 is loaded into the RAM 115, it can be used by the processor 112. In case of large software applications 136, processor 112 loads various portions of program modules into RAM 115 as needed.

30 The Basic Input/Output System (BIOS) 117 for the computer 100 is stored in ROM 114 and is loaded into RAM 115 upon booting. Those skilled in the art will recognize that the

BIOS 117 is a set of basic executable routines that have conventionally helped to transfer information between the computing resources within the computer 100. Operating system 135 or other software applications 136 use these low-level service routines.

In one embodiment computer 100 includes a registry (not shown) which is a system database that holds configuration information for computer 100. For example, Windows® 95 and Windows® NT by Microsoft maintain the registry in two hidden files, called USER.DAT and SYSTEM.DAT, located on a permanent storage device such as an internal disk.

Methods and Systems of the Invention

Figure 2 is a block diagram illustrating one embodiment of a computing environment 200 that provides seamless access to remote storage areas. In the illustrated embodiment, client computers 205 are communicatively coupled to remote storage network 220 via storage servers 210 and global computer network 215 such as the Internet. Storage network 220 represents one or more interconnected storage devices, such as a RAID, for storing data files. The storage network 220 defines a pool of virtual storage areas 225 that can be individually assignable to different users. Initially, a user accesses storage servers 210 via global computer network and requests a virtual storage area 225. For each request, storage server 210 prompts the user for information such as a username, password, billing address, etc. Upon approval, storage network 220 allocates a storage area 225 to the user such that, as described in detail below, the user can seamlessly access the corresponding virtual storage area via client computers 205. Metadata database 222 stores metadata associated with the data files by the user. Storage server 210 and client computers 205 represent computing devices such as computer 100 described in Figure 1.

According to various embodiments of the invention, a set of communication software applications execute on storage servers 210 and provide a set of interfaces such that client computers 205 can access virtual storage areas 225 using any one of several different methods. Each method offers different advantages in control, performance and ease of use. For example, in one embodiment, access to the virtual storage areas 225 is fully integrated with operating system 135 executing on each client computer 205. In this embodiment, the user can directly access a virtual storage area 225 via the operating system's user interface. Software applications 136 executing on client computer 205 can access the corresponding

virtual storage area 225 through operating system 135 by calling standard file management routines provided by operating system 135. In this embodiment, operating system 135 packetizes data files and metadata received from the management routines and communicates the data to storage servers 210 via network 215. In this embodiment, operating system 135 handles all communication with storage servers 210 such that virtual storage area 225 can be accessed as if it were local to client computer 205.

In another embodiment, the user can access virtual storage area 225 by invoking a communications application 136 such as a web browser or an FTP utility. In this embodiment, the communications application 136 handles communication across network 215. The underlying communications between client computers 205 and storage servers 210 can follow a variety of protocols. Typically, the communications use TCP/IP as a base protocol and additionally use the HTTP protocol, the FTP protocol or even a proprietary data-backup protocol.

The following sub-sections describe the various techniques by which, according to the invention, client computers 205 access virtual storage areas 225 via storage servers 210 and global computer network 215.

Access via the Operating System

According to the invention, in one embodiment, the user can access virtual storage area 225 via operating system 135 in the same manner by which the user accesses local storage devices 120, 121 and 122. In this embodiment, the user interacts with the user interface presented by operating system 135 and manages virtual storage area 225 as if it were a local storage volume. For example, by interacting with operating system 135 the user can perform all common file management tasks including copying files between hard disk 120 and remote storage area 225, as well as renaming and deleting data files. The user is able to browse directory structures within virtual storage area 225 as if it were a local storage volume. Indeed, in some embodiments, virtual storage area 225 may be represented as an icon on the desktop of the user's computer.

Figure 3 illustrates window 300 as displayed by operating system 135 for accessing a virtual storage area 225. In the illustrated embodiment, operating system 135 is the Macintosh® operating system from Apple Computer, Inc. Window 300 presented by

operating system 135 lists the directories and files of virtual storage area 225 as if the directories and files were local. The user is able to browse the directories within virtual storage area 225 and perform the file management operations supported by operating system 135 such as copying, renaming, moving and deleting files and directories. Similarly, applications 136 executing on client computer 205 access virtual storage area 225 through standard file management routines provided by an application programming interface (API) of operating system 135. In one embodiment, operating system 135 of client computers 205 is the Macintosh operating system, such that the API includes the Apple File Services (AFS), and storage servers 225 are support accessing remote data files within storage area virtual 225 via the Apple Filing Protocol (AFP) services over TCP/IP.

Figure 4 illustrates another embodiment in which operating system 135 is the Windows® operating system from Microsoft. Here, operating system 135 displays window 400 that presents an integrated list of those files and directories that are available locally as well as those available from virtual storage area 225. In this embodiment, operating system 135 incorporates the SMB protocol or the WebDAV protocol in order to provide seamless access to virtual storage area 225.

One particular advantage of this embodiment is that software applications 136 executing on computer 100 can access virtual storage area 225 without modification. Because operating system 135 provides seamless access to virtual storage area 225 via conventional file management routines, software applications 136 treat virtual storage area 225 as if it were a local storage volume. Another advantage is that the user can write scripts for operating system 135, such as an OSAX AppleScript for the Macintosh® operating system, in order to control access to virtual storage area 225. Figure 5 illustrates an operating system script 500 written in AppleScript that automatically copies a file from a local directory to a similarly named remote directory located within virtual storage area 225.

According to one aspect of the invention, the user can associate metadata with each data file as the data file is uploaded to remote storage areas 225 in order to facilitate the quick cataloging and quick retrieval of the data file. Operating system 135 captures the metadata from the user via an input form and communicates the metadata to storage servers 210.

Direct Access From a Communication Software Application

In some embodiments, the user can invoke conventional communication applications and utilities such as a web browser, an FTP utility, a data storage backup utility, or even a conventional email software application, to access virtual storage area. During this method of access, the communication software application 136 handles all communications with storage servers 210. The file management routines of operating system 135 are not invoked. In addition, the communication software 136 is responsible for capturing the metadata from the user and communicating the metadata to storage servers 210 for storage within metadata database 222.

Figure 6 illustrates the user accessing one of the virtual storage areas 225 via a conventional web browser executing on client computer 205. The web browser displays window 600 that lists each directory within virtual storage area 225. Storage servers 210 maintain a set of image icons for representing the stored data file according to file and creator type information or file extension. Storage servers 210 select and display an appropriate icon as a function of the file and creator information stored within virtual storage area 225. Window 600, therefore, closely resembles windows 300 and 400 as displayed by operating system 135. Using the browser, the user can browse the directories within virtual storage area 225 and can perform many common file management operations including uploading, downloading and deleting files, as well as creating and removing directories.

Figure 7 illustrates the user accessing virtual storage area 225 using a conventional electronic mail software application 136. Electronic mail software application 136 displays window 700 by which the user creates an email message having user information 702 and target data file information 710. The user addresses the email to his or her email address at storage servers 210. For uploading files, as displayed in window 700, the user attaches one or more data files 705. Upon receiving the email, storage server 210 parses the user information 702 and the target data file information 710 in order to extract one or more target directories. Thus, user information 702 may take the form of a destination electronic mail address, while target data file information 710 may take the form of text in the subject field of an electronic mail message. Storage server 210 stores attached files 705 to the target directory within the appropriate virtual storage area 225 assigned to the user. In this manner, the user is able to email one or more files directly into a specified directory within virtual

storage area 225 from any remote location via network 215. In an alternative embodiment, the user does not attach a data file but includes within the electronic mail message a universal resource locator (URL) that indicates where storage server 210 can retrieve the data file to be stored. In order to retrieve one or more data files from storage server 210, the user sends a similar email requesting one or more target data files. In response, storage server 210 sends the user a response electronic mail message with the requested data files attached.

Remote Processing via Centrally Hosted Applications

The above sections described how, according to the invention, client computers 205 access virtual storage area 225 for remote data storage. The invention, however, is not limited to data storage. According to one aspect of the invention, client computers 205 communicate data files to storage servers 210 for remote processing by one or more applications executing on storage servers 210. For example, the communicated data files may be electronic images such that applications executing on storage servers 210 perform color correction, color proofing and other prepress operations. Other examples include data conversion, compression, decompression, encryption and decryption. As such, a user is able to submit data files for remote processing by simply instructing operating system 135 to copy the data files to a designated directory within virtual storage area 225. As such, this operation can be performed by simply dragging and dropping the file onto the appropriate window presented by operating system 135.

Allocation of a Remote Storage Area

Figure 8 is a flow chart 800 that provides a high-level overview of one mode of operation in which storage network 220 allocates remote storage areas 225. In block 803, users access storage servers 210 via global computer network 215 and request a corresponding virtual storage area 225. For each request, storage server 210 prompts the user for information such as a username, password, billing address (block 805). Upon approval, storage network 220 automatically allocates a corresponding storage area 225 for each user and stores the respective user information in metadata database 222 (block 807). After a virtual storage area 25 has been allocated, storage network 220 informs each user of any necessary access information, such as a password, so that the user can access the respective

storage area 225 via the many access interfaces described above (block 809).

Various embodiments have been described of a method and system for a data storage system that provides seamless access to remote virtual data storage areas via a global computer network. Access to the virtual storage area is fully integrated with the operating system executing on the client computer such that the user, or software applications
5 executing on the computer, can seamlessly access the corresponding virtual storage area using standard file management routines provided by the operating system. In addition, the invention provides a wide range of other remote access methods to the virtual data storage areas including access via electronic mail. This application is intended to cover any
10 adaptation or variation of the present invention. It is intended that this invention be limited only by the claims and equivalents thereof.